audio system providing for filter coefficient copying

The present invention relates to a system for suppressing audio distortion, comprising a circuit arrangement of:

- echo cancelling means coupled between an audio output and a distorted desired audio sensing microphone array, and
- a filter arrangement coupled to the echo cancelling means and/or the microphone array.

The present invention also relates to a mirrored circuit arrangement for application in the system and to a method of suppressing audio distortion.

Such a system is known from WO 97/45995. The known audio system comprises an adaptive echo cancelling filter for removing echoes emanating between a systems' loudspeaker output and a microphone. The known system has a filter arrangement coupled to the echo cancelling filter and the microphone for spectrally suppressing echo components in the microphone signal that were not removed by the echo cancelling filter. One microphone senses a desired audio signal, while the other microphones only receive interfering distortions of the desired signal. The system may have a filter arrangement coupled to the echo cancelling means and/or the microphone array for spectrally suppressing distortion in the form of additional audio noise interference.

It is a disadvantage of the known system that it can not effectively be used for also reducing reverberant distortions in a desired audio signal sensed by a microphone array.

Therefore it is an object of the present invention to provide an improved system and filter arrangement therein for also suppressing echo distortion in the form of echo tail part reverberation in an audio signal sensed by a microphone array.

Thereto in the system according to the invention the filter arrangement includes filter coefficients representing at least a part of the audio distortion, and the system comprises an at least partly mirrored circuit arrangement for copying thereby simulated audio

distortion representative filter coefficient values into the filter coefficients of said filter arrangement.

It is an advantage of the system and circuit arrangement according to the present invention that by copying simulated filter coefficients which are representative of the audio distortion into the filter coefficients of the filter arrangement the suppression of the audio distortion is made speech signal –in general desired signal- independent. The coefficient values represent the correlation properties of the reverberant tail part(s) of the sound field and can also be used for actually filtering the desired signal in filter arrangement for suppressing reverberation therein, irrespective whether desired speech is output or not. So in particular the presence or absence of speech no longer distorts the distortion cancellation. Even if speech is present distortion cancellation of the spatial correlation sensitive forms of distortion, in particular but not exclusively, reverberation types of distortion, can be effected efficiently by means of the system according to the present invention.

An embodiment of the system according to the invention allowing design flexibility is characterised in that the filter arrangement includes a beamformer.

Most often a combination of filter, sum and delay elements is comprised in the filter arrangement to form the so called Generalised Sidelobe Canceller. Advantageously an additional delay element may be added to the beamformers for further improving the performance of the system according to the invention.

Advantageously another simple embodiment of the system according to the invention is characterised in that the system comprises coefficient value copying means between the circuit arrangement and the at least partly mirrored circuit arrangement.

A further embodiment of the system according to the invention is characterised in that the filter arrangement is arranged to be adaptive to the reverberation distortion and/or the desired audio signal sensed by the microphone array.

In that case the filter coefficients can be updated, to include a dynamic aspect in the cancelling of varying correlation sensitive forms of distortion, such as for example reverberation, instead of representing a more or less fixed model of the room. Now such distortion can also be suppressed in relation to the respective varying positions and directions of the array microphones.

A still further embodiment of the system according to the invention is characterised in that the system is arranged for updating the mirrored filter coefficients.

Advantageously in this further embodiment the mirrored filter coefficients may be updated continuously, whenever necessary or during a training session. This is

achieved because the mirrored filter coefficients are created by means of an artificial input signal representing reverberation.

An elaboration of the system according to the invention is characterised in that each microphone of the microphone array has its individual echo cancelling means.

By applying individualised echo cancelling means for each microphone of the array any separate direct echoes and reflections, and at least a part of the reverberating tail are cancelled individually as much as possible, while remaining distortion is dealt with by the filter arrangement and/or the output echo cancelling means.

At present the system according to the invention will be elucidated further together with its additional advantages, while reference is being made to the appended drawing, wherein similar components are being referred to by means of the same reference numerals. In the drawing:

Fig. 1 shows a overview including possible embodiments of the system according to the invention;

Fig. 2 shows the direct signal, the early reflections and the later arising reverberation tail of a typical room impulse response as a function of time; and

Fig. 3 shows a filter arrangement embodiment in the form of a generalised sidelobe canceller having an array of three microphones for application in an extension of the system of Fig. 1.

The upper part of fig. 1 shows a system 1, which is suited for suppressing audio distortion in a desired signal. The system as shown has a loudspeaker 2 and a microphone array 3 comprising two microphones 3-1, 3-2. An audio output signal on output 4 is reproduced by the loudspeaker 2. A near end source (not shown) generates desired speech, which is received by the array 3 as a desired speech signal. In addition the array 3 senses -as clarified in Fig. 2- as part of different kinds of distortions apart from noise, a direct signal from the loudspeaker 2 to the array 3, echoes in the form of early –first part- reflections and after some exponential decay later –second part- reflections in the form of so called reverberation shown as a reverberating tail of a typical room impulse response as a function of time. Each microphone 3-1, 3-2 may have its associated echo canceller g_1, and g_2 respectively coupled between the audio output 4 and the distorted desired audio sensing

microphone array 3. If at all possible hardware and/or software parts of the echo cancelling means g_i ($i = 1, 2$ for two microphones) may be used in common in order to save costs. Each of the echo cancellers g_i simulate the path from the loudspeaker 2 to a respective microphone 3 in order to cancel the effects of at least the direct signal and the early reflections, that is the first part of the echo. The technique accomplishing that is for example known from WO 97/45995, whose disclosure is incorporated herein by reference thereto. The respective echo cancelling means may be implemented in various ways, such as with Least Mean Squares (LMS), Recursive Least Squares or Frequency Domain Adaptive Filter using Block LMS techniques.

The respective echo cancelling means g_i are coupled to two microphones 3-1, 3-2 of the array 3 through schematically shown subtractors 5-1, and 5-2 each having outputs 6-1, 6-2. These subtractor outputs 6-1 and 6-2 carry respective echo cancelled signals.

The system has a filter arrangement 7, which may include a beamformer 7B, which is coupled through the subtractors 5 to the echo cancelling means g_i and/or to the microphone array 3. The beamformer 7B, which is included in a generally called Generalised Sidelobe Canceller, is capable of defining and controlling an audio microphone sensitivity lob or curve. Given the in this case two beamformer input signals on the subtractor outputs 6-1, 6-2, these signals comprise the desired audio/sound/speech signal and a reverberation signal originating from the reverberating tail. The beamformer 7B is capable of discriminating the reverberation signal by deriving a primary signal z including the desired signal and a reference signal x which includes the reverberation. It does this here by filtering in filters f_1 and f_2, as shown, and then summing in summing device 9-1 the filters f_i outputs to reveal the primary signal. This way the echo cancelled microphone signals u_1 and u_2 are added such that remaining direct signals and early reflections of the desired audio are coherently summed, which increases the beamformers performance. Furthermore it does this here by filtering the echo cancelled microphone signals in blocking filters b_1 and b_2 and then by summing in device 9-2 the filters' outputs to reveal a reverberation representing reference signal x. The reference signal x virtually contains no desired signal components. The filters b_i together B, are called the blocking matrix. The filters f_i and b_i carry the directional, that is the desired sources dependent information.

In the case as shown in Fig. 3 the beamformer 7B has one delay element 8 coupled to output 10 of device 9-1 followed by a summing device 9-3. The delay element 8 provides a non causal part to the beamformers' impulse response which appeared to improve its performance. The reference signal x is fed to an adaptive filter, indicated w in Fig. 1,

whose output signal is fed to an inverting input 11 of device 9-3. The filter w of the filter arrangement 7 comprises the filter coefficients which represent or contain a measure for the reverberation –second part- distortion in the desired audio sensed by the microphone array 3. The summing device 9-3 also has a summed or beamformer output S used to adapt the filter coefficients in the adaptive filter w of the thus adaptive filter arrangement 7, such that their coefficient values represent the varying reverberation distortion. In a non adaptive embodiment the filter coefficients would be fixed to then cancel a then presumed fixed reverberation tail.

Because the filtered reverberation or reference signal on inverting input 11 is subtracted from the primary signal in summing device 9-3 its signal on the summed output S only contains the desired signal, with the reverberating tail being cancelled.

Fig. 3 shows an embodiment of a filter arrangement 7 having an array of three microphones 3-1, 3-2, 3-3. Essentially a plurality of microphones is possible. However above outlined principles remain the same. Block matrices may be grouped into one block B. Different reference signals x_1 and x_2 may be fed to the filter 7A, here comprising generally adaptive individualised filters w_1 and w_2. At wish delay elements Δ may be divided up in front or after the filters f_1, f_2, and f_3 coupled to the respective three microphones 3. Separate delay elements Δ could be included in the respective branches from possibly each of the microphones to summing device 9-1 to account for expected individual delays between loudspeaker 2 and microphone 3.

If the system 1 does not start up by itself, due to absence of any far end signal a loudspeaker signal could be generated, e.g. a noise sequence or some kind of start up tune.

The system explained above can for example be used in hands-free communication systems, such as hands-free speakerphones, voice controlled systems for example in home or for medical applications, congress systems, dictation system or the like.

Although the above description of the figures is related to a filter arrangement embodied by a beamformer 7 it should be noted that the above also holds for any filter arrangement 7 which is aimed at suppressing audio distortion in general in the system 1 as elucidated in the above.

Reverberation is a form of distortion, whose cancellation is strongly dependent on the spatial correlation properties of the microphone array formation and the room concerned. The spatial correlation of the sound field determines the mutual correlation between the microphone signals, which in turn is the quantity which determines the design and performance of the filter arrangement and beamformer 7. These spatial correlation

properties depend on several properties such as room geometry, wall absorption, position, direction and spacing of the microphones of the array 3 in a room. Advantageously the system 1 does not require some kind of model for these properties. When however these properties and/or modelled distortion echo cancelling properties or coefficients in the system 1 change significantly then substantial adaptation thereof is required, which takes a considerable amount of time. In general this time is not always available having the consequence of audible and disturbing effects during in particular shortly after presence or after absence of especially speech. The above is of primary importance in relation to spatial correlation dependent forms of distortion, such as reverberation. The remainder of this description will therefore be directed to cancellation of reverberation by means of the system having a filter arrangement as shown in figure 1.

The lower part of fig. 1 shows a circuit arrangement 7' whose components are at least partly similar to –or mirrored relative- to the echo cancelling means g and the beamformer filter 7. It comprises the beamformer 7' which may have parts similar to beamformer 7, and the means 9' here having g_1' and g_2' as echo filters, whose coefficients may be influenced or set. The respective echo filter means g'_i generally for i = 1, 2, .. are simple filters normally each have fewer coefficients than the number of coefficients in the corresponding echo canceller means g_i at least in the case of reverberation cancellation, because then only the reverberation tail part of the simulated impulse response has to be copied by schematically shown copying means C1, C2 and C3 from the canceller means g_i into the filter means g_i'. Similarly the filter characteristics or coefficient values of the filters f_1, f_2, b_1, and b_2 are copied into the respective filters f_1', f_2', b_1', and b_2' of beamfilter 7'. Now an auxiliary input signal is applied on input terminal 12 of the mirrored circuit arrangement g', 7' such that the adaptive filter w' in 7' which is coupled to output S' is capable of minimizing its output signal on output S'. The auxiliary input signal on terminal 12 may for example be a stationary white noise signal or any other signal depending on the specific type of distortion to be cancelled. The thus simulated audio distortion representative filter coefficient values of the filter w' are copied into the filter coefficients of the filter w in filter arrangement 7. These values represent the correlation properties of the reverberant tail part(s) of the sound field and can also be used for actually filtering the desired signal in filter arrangement 7 for suppressing reverberation therein, irrespective whether desired speech is output or not. In order to adapt or update the filter 7' beamformer coefficients, only the reverberant behaviour of the room needs to taken into account. Thereto the desired audio

source is not required, as any source in the room would do that job. As far as the filter w receives its coefficients from the adaptive filter w' the filter w does not have to be adaptive.

audio system providing for filter coefficient copying

The present invention relates to a system for suppressing audio distortion, comprising a circuit arrangement of:

- echo cancelling means coupled between an audio output and a distorted desired audio sensing microphone array, and
- a filter arrangement coupled to the echo cancelling means and/or the microphone array.

The present invention also relates to a mirrored circuit arrangement for application in the system and to a method of suppressing audio distortion.

Such a system is known from WO 97/45995. The known audio system comprises an adaptive echo cancelling filter for removing echoes emanating between a systems' loudspeaker output and a microphone. The known system has a filter arrangement coupled to the echo cancelling filter and the microphone for spectrally suppressing echo components in the microphone signal that were not removed by the echo cancelling filter. One microphone senses a desired audio signal, while the other microphones only receive interfering distortions of the desired signal. The system may have a filter arrangement coupled to the echo cancelling means and/or the microphone array for spectrally suppressing distortion in the form of additional audio noise interference.

It is a disadvantage of the known system that it can not effectively be used for also reducing reverberant distortions in a desired audio signal sensed by a microphone array.

Therefore it is an object of the present invention to provide an improved system and filter arrangement therein for also suppressing echo distortion in the form of echo tail part reverberation in an audio signal sensed by a microphone array.

Thereto in the system according to the invention the filter arrangement includes filter coefficients representing at least a part of the audio distortion, and the system comprises an at least partly mirrored circuit arrangement for copying thereby simulated audio

distortion representative filter coefficient values into the filter coefficients of said filter arrangement.

It is an advantage of the system and circuit arrangement according to the present invention that by copying simulated filter coefficients which are representative of the audio distortion into the filter coefficients of the filter arrangement the suppression of the audio distortion is made speech signal –in general desired signal- independent. The coefficient values represent the correlation properties of the reverberant tail part(s) of the sound field and can also be used for actually filtering the desired signal in filter arrangement for suppressing reverberation therein, irrespective whether desired speech is output or not. So in particular the presence or absence of speech no longer distorts the distortion cancellation. Even if speech is present distortion cancellation of the spatial correlation sensitive forms of distortion, in particular but not exclusively, reverberation types of distortion, can be effected efficiently by means of the system according to the present invention.

An embodiment of the system according to the invention allowing design flexibility is characterised in that the filter arrangement includes a beamformer.

Most often a combination of filter, sum and delay elements is comprised in the filter arrangement to form the so called Generalised Sidelobe Canceller. Advantageously an additional delay element may be added to the beamformers for further improving the performance of the system according to the invention.

Advantageously another simple embodiment of the system according to the invention is characterised in that the system comprises coefficient value copying means between the circuit arrangement and the at least partly mirrored circuit arrangement.

A further embodiment of the system according to the invention is characterised in that the filter arrangement is arranged to be adaptive to the reverberation distortion and/or the desired audio signal sensed by the microphone array.

In that case the filter coefficients can be updated, to include a dynamic aspect in the cancelling of varying correlation sensitive forms of distortion, such as for example reverberation, instead of representing a more or less fixed model of the room. Now such distortion can also be suppressed in relation to the respective varying positions and directions of the array microphones.

A still further embodiment of the system according to the invention is characterised in that the system is arranged for updating the mirrored filter coefficients.

Advantageously in this further embodiment the mirrored filter coefficients may be updated continuously, whenever necessary or during a training session. This is

achieved because the mirrored filter coefficients are created by means of an artificial input signal representing reverberation.

An elaboration of the system according to the invention is characterised in that each microphone of the microphone array has its individual echo cancelling means.

By applying individualised echo cancelling means for each microphone of the array any separate direct echoes and reflections, and at least a part of the reverberating tail are cancelled individually as much as possible, while remaining distortion is dealt with by the filter arrangement and/or the output echo cancelling means.

At present the system according to the invention will be elucidated further together with its additional advantages, while reference is being made to the appended drawing, wherein similar components are being referred to by means of the same reference numerals. In the drawing:

Fig. 1 shows a overview including possible embodiments of the system according to the invention;

Fig. 2 shows the direct signal, the early reflections and the later arising reverberation tail of a typical room impulse response as a function of time; and

Fig. 3 shows a filter arrangement embodiment in the form of a generalised sidelobe canceller having an array of three microphones for application in an extension of the system of Fig. 1.

The upper part of fig. 1 shows a system 1, which is suited for suppressing audio distortion in a desired signal. The system as shown has a loudspeaker 2 and a microphone array 3 comprising two microphones 3-1, 3-2. An audio output signal on output 4 is reproduced by the loudspeaker 2. A near end source (not shown) generates desired speech, which is received by the array 3 as a desired speech signal. In addition the array 3 senses -as clarified in Fig. 2- as part of different kinds of distortions apart from noise, a direct signal from the loudspeaker 2 to the array 3, echoes in the form of early –first part- reflections and after some exponential decay later –second part- reflections in the form of so called reverberation shown as a reverberating tail of a typical room impulse response as a function of time. Each microphone 3-1, 3-2 may have its associated echo canceller g_1, and g_2 respectively coupled between the audio output 4 and the distorted desired audio sensing

microphone array 3. If at all possible hardware and/or software parts of the echo cancelling means g_i ($i = 1, 2$ for two microphones) may be used in common in order to save costs. Each of the echo cancellers g_i simulate the path from the loudspeaker 2 to a respective microphone 3 in order to cancel the effects of at least the direct signal and the early reflections, that is the first part of the echo. The technique accomplishing that is for example known from WO 97/45995, whose disclosure is incorporated herein by reference thereto. The respective echo cancelling means may be implemented in various ways, such as with Least Mean Squares (LMS), Recursive Least Squares or Frequency Domain Adaptive Filter using Block LMS techniques.

The respective echo cancelling means g_i are coupled to two microphones 3-1, 3-2 of the array 3 through schematically shown subtractors 5-1, and 5-2 each having outputs 6-1, 6-2. These subtractor outputs 6-1 and 6-2 carry respective echo cancelled signals.

The system has a filter arrangement 7, which may include a beamformer 7B, which is coupled through the subtractors 5 to the echo cancelling means g_i and/or to the microphone array 3. The beamformer 7B, which is included in a generally called Generalised Sidelobe Canceller, is capable of defining and controlling an audio microphone sensitivity lob or curve. Given the in this case two beamformer input signals on the subtractor outputs 6-1, 6-2, these signals comprise the desired audio/sound/speech signal and a reverberation signal originating from the reverberating tail. The beamformer 7B is capable of discriminating the reverberation signal by deriving a primary signal z including the desired signal and a reference signal x which includes the reverberation. It does this here by filtering in filters f_1 and f_2, as shown, and then summing in summing device 9-1 the filters f_i outputs to reveal the primary signal. This way the echo cancelled microphone signals u_1 and u_2 are added such that remaining direct signals and early reflections of the desired audio are coherently summed, which increases the beamformers performance. Furthermore it does this here by filtering the echo cancelled microphone signals in blocking filters b_1 and b_2 and then by summing in device 9-2 the filters' outputs to reveal a reverberation representing reference signal x. The reference signal x virtually contains no desired signal components. The filters b_i together B, are called the blocking matrix. The filters f_i and b_i carry the directional, that is the desired sources dependent information.

In the case as shown in Fig. 3 the beamformer 7B has one delay element 8 coupled to output 10 of device 9-1 followed by a summing device 9-3. The delay element 8 provides a non causal part to the beamformers' impulse response which appeared to improve its performance. The reference signal x is fed to an adaptive filter, indicated w in Fig. 1,

whose output signal is fed to an inverting input 11 of device 9-3. The filter w of the filter arrangement 7 comprises the filter coefficients which represent or contain a measure for the reverberation –second part- distortion in the desired audio sensed by the microphone array 3. The summing device 9-3 also has a summed or beamformer output S used to adapt the filter coefficients in the adaptive filter w of the thus adaptive filter arrangement 7, such that their coefficient values represent the varying reverberation distortion. In a non adaptive embodiment the filter coefficients would be fixed to then cancel a then presumed fixed reverberation tail.

Because the filtered reverberation or reference signal on inverting input 11 is subtracted from the primary signal in summing device 9-3 its signal on the summed output S only contains the desired signal, with the reverberating tail being cancelled.

Fig. 3 shows an embodiment of a filter arrangement 7 having an array of three microphones 3-1, 3-2, 3-3. Essentially a plurality of microphones is possible. However above outlined principles remain the same. Block matrices may be grouped into one block B. Different reference signals x_1 and x_2 may be fed to the filter 7A, here comprising generally adaptive individualised filters w_1 and w_2. At wish delay elements Δ may be divided up in front or after the filters f_1, f_2, and f_3 coupled to the respective three microphones 3. Separate delay elements Δ could be included in the respective branches from possibly each of the microphones to summing device 9-1 to account for expected individual delays between loudspeaker 2 and microphone 3.

If the system 1 does not start up by itself, due to absence of any far end signal a loudspeaker signal could be generated, e.g. a noise sequence or some kind of start up tune.

The system explained above can for example be used in hands-free communication systems, such as hands-free speakerphones, voice controlled systems for example in home or for medical applications, congress systems, dictation system or the like.

Although the above description of the figures is related to a filter arrangement embodied by a beamformer 7 it should be noted that the above also holds for any filter arrangement 7 which is aimed at suppressing audio distortion in general in the system 1 as elucidated in the above.

Reverberation is a form of distortion, whose cancellation is strongly dependent on the spatial correlation properties of the microphone array formation and the room concerned. The spatial correlation of the sound field determines the mutual correlation between the microphone signals, which in turn is the quantity which determines the design and performance of the filter arrangement and beamformer 7. These spatial correlation

properties depend on several properties such as room geometry, wall absorption, position, direction and spacing of the microphones of the array 3 in a room. Advantageously the system 1 does not require some kind of model for these properties. When however these properties and/or modelled distortion echo cancelling properties or coefficients in the system 1 change significantly then substantial adaptation thereof is required, which takes a considerable amount of time. In general this time is not always available having the consequence of audible and disturbing effects during in particular shortly after presence or after absence of especially speech. The above is of primary importance in relation to spatial correlation dependent forms of distortion, such as reverberation. The remainder of this description will therefore be directed to cancellation of reverberation by means of the system having a filter arrangement as shown in figure 1.

The lower part of fig. 1 shows a circuit arrangement 7' whose components are at least partly similar to –or mirrored relative- to the echo cancelling means g and the beamformer filter 7. It comprises the beamformer 7' which may have parts similar to beamformer 7, and the means 9' here having g_1' and g_2' as echo filters, whose coefficients may be influenced or set. The respective echo filter means g'_i generally for i = 1, 2, .. are simple filters normally each have fewer coefficients than the number of coefficients in the corresponding echo canceller means g_i at least in the case of reverberation cancellation, because then only the reverberation tail part of the simulated impulse response has to be copied by schematically shown copying means C1, C2 and C3 from the canceller means g_i into the filter means g_i'. Similarly the filter characteristics or coefficient values of the filters f_1, f_2, b_1, and b_2 are copied into the respective filters f_1', f_2', b_1', and b_2' of beamfilter 7'. Now an auxiliary input signal is applied on input terminal 12 of the mirrored circuit arrangement g', 7' such that the adaptive filter w' in 7' which is coupled to output S' is capable of minimizing its output signal on output S'. The auxiliary input signal on terminal 12 may for example be a stationary white noise signal or any other signal depending on the specific type of distortion to be cancelled. The thus simulated audio distortion representative filter coefficient values of the filter w' are copied into the filter coefficients of the filter w in filter arrangement 7. These values represent the correlation properties of the reverberant tail part(s) of the sound field and can also be used for actually filtering the desired signal in filter arrangement 7 for suppressing reverberation therein, irrespective whether desired speech is output or not. In order to adapt or update the filter 7' beamformer coefficients, only the reverberant behaviour of the room needs to taken into account. Thereto the desired audio

source is not required, as any source in the room would do that job. As far as the filter w receives its coefficients from the adaptive filter w' the filter w does not have to be adaptive.